Exploring Student Growth in Missouri

Changing Conversations about Education

Damian W. Betebenner

National Center for the Improvement of Educational Assessment
Dover, New Hampshire

Measuring Teacher and Leader Effect Work Group
Columbia, Missouri March 25th, 2010

About Me

Dr. Damian Betebenner, PhD

Senior Associate at the Center for Assessment (NCIEA). Developed student growth percentiles and percentile growth trajectories to help states and educational associations employ student growth in decision making [Betebenner, 2008, Betebenner, 2009]. Currently refining and sharing these techniques with other states including Colorado, Massachusetts, Arizona, Indiana, as well as at least 15 other states in various stages of investigation/adoption.

Today's Topics

- Questions that a growth model can/should address.
- Student growth percentiles.
- Student growth versus value-added
- Changing conversations around education using growth data.

Data Data Everywhere

Enhanced data acquisition and management has enabled:

- Historical records of student achievement
- Historical records of student demographics, teachers, schools, educational programs, . . .
- Stakeholder interest in an examination of this longitudinal data

Interest in examining student achievement over time (student growth) derives from data availability.

The attractiveness of growth

What is growth and why measure it?

- Student learning is a central goal of education.
- Assessments of student achievement provide evidence of the current status of student knowledge and understanding.
- Learning is demonstrated by growth in student achievement from one point in time to another point in time–not by status at either point time alone.

Accountability & Growth

Growth versus Status

- Enthusiasm for growth in accountability stems from the belief that growth and teacher/school quality are more closely related than status.
- Enthusiasm for growth also stems from its potential diagnostic uses.
- How do we judge the use of growth related measures within an accountability system?

Guiding principals

John Tukey

It is better to have an approximate answer to the right question than a precise answer to the wrong question.

George E. P. Box

All models are wrong but some are useful.

Start at the end and work backward

The best growth model answers the questions of greatest relevance to stakeholders.

Guiding questions

Growth models address specific questions

- Different growth analysis techniques are good at answering different questions.
- It is critical to understand these different questions.
- Different questions lead to different conversations which lead to different uses and outcomes.

Changing conversations about education

- Starting with the right questions simplifies development and motivating the proper use of the growth model results.
- It's all about the conversations you want to have. The questions set the table for those conversations.

What are the relevant questions for parents?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

Parent Questions

- Did my child make a year's worth of progress in a year?
- Is my child growing appropriately toward meeting state standards?
- Is my child growing as much in Math as Reading?
- Did my child grow as much this year as last year?

What are the relevant questions for teachers?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

Teacher Questions

- Did my students make a year's worth of progress in a year?
- Did my students grow appropriately toward meeting state standards?
- How close are my students to becoming Proficient?
- Are there students with unusually low growth who need special attention?

What are the relevant questions for administrators?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

Administrator Questions

- Did the students in our district/school make a year's worth of progress in all content areas?
- Are our students growing appropriately toward meeting state standards?
- Does this school/program show as much growth as that one?
- Can I measure student growth even for students who do not change proficiency categories?
- Can I pool together results from different grades to draw summary conclusions?

The Colorado Growth Model

How much growth did a student make? Is it enough?

- An important first step in the development of the Colorado Growth Model was to separate the description of growth from discussions of responsibility (i.e., accountability).
- Incorporating growth into accountability follows more easily from the viable description of growth.
- A viable description of growth facilitates stakeholder engagement and investigations of responsibility for good/bad growth.
- This in turn leads to greater stakeholder support for particular forms of accountability.

Describing student growth

- Measuring student growth, even with a vertical scale, is not a simple task.
- Some believe a vertical scale simplifies the task of measuring student growth.
- Even with an interval (or ratio) scale, growth is not easy to interpret. Consider, for example, height.
  - A child might grow 4 inches between ages 3 and 4.
  - 4 inches is a well understood quantity.
  - The 4 inch increase becomes really meaningful only when understood alongside the growth of other 3 to 4 year olds.
- Student growth percentiles were developed to provide a normative context for describing student growth.

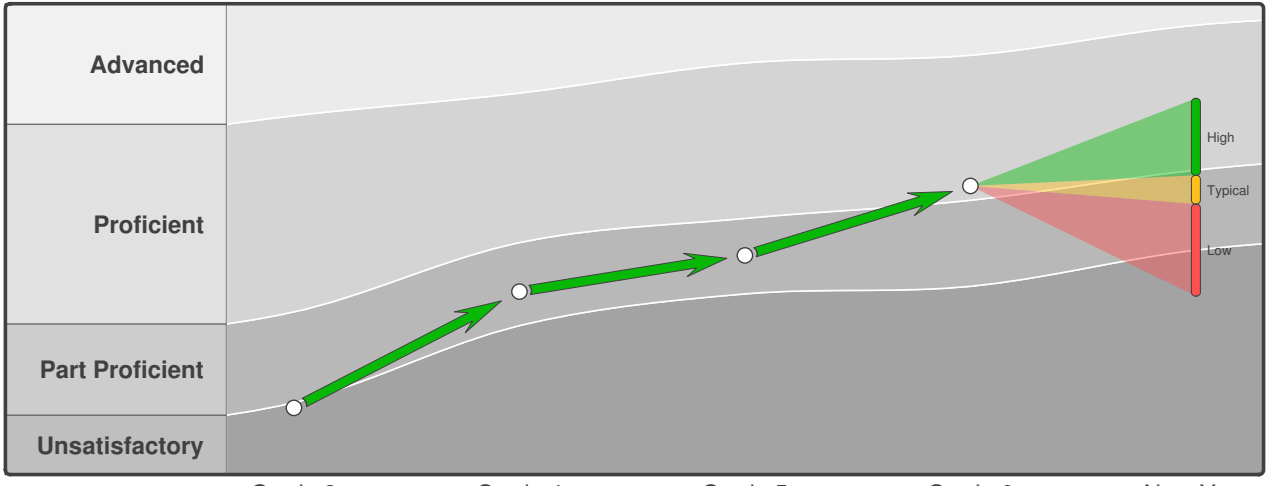
Student Growth Percentiles

Should we be surprised with a child's current achievement given their prior achievement?

- Student growth percentiles answer this question.
- Consider a low achieving student with 90th percentile growth and a high achieving student with 10th percentile growth.
  - The low achieving student grew at a rate exceeding 90 percent of similar students.
  - The high achieving student grew at a rate exceeding just 10 percent of similar students.
  - The low achiever's growth is more *exemplary* (probabilistically) than the high achiever's.
- Judgments about the adequacy of student growth require external criteria.

Combining norms and standards

- Growth adequacy is determined by whether a student's growth is sufficient to reach/maintain desired achievement levels (e.g., proficiency).
- A next step for the Archdiocese is to consider establishing performance standards (ideally linked to content standards) that can be used to qualify student achievement.
- With established performance standards, percentile growth projections/trajectories are calculated for each student to reach/maintain desired levels of achievement.



Reading

Achievement

CSAP Reading Scale Score

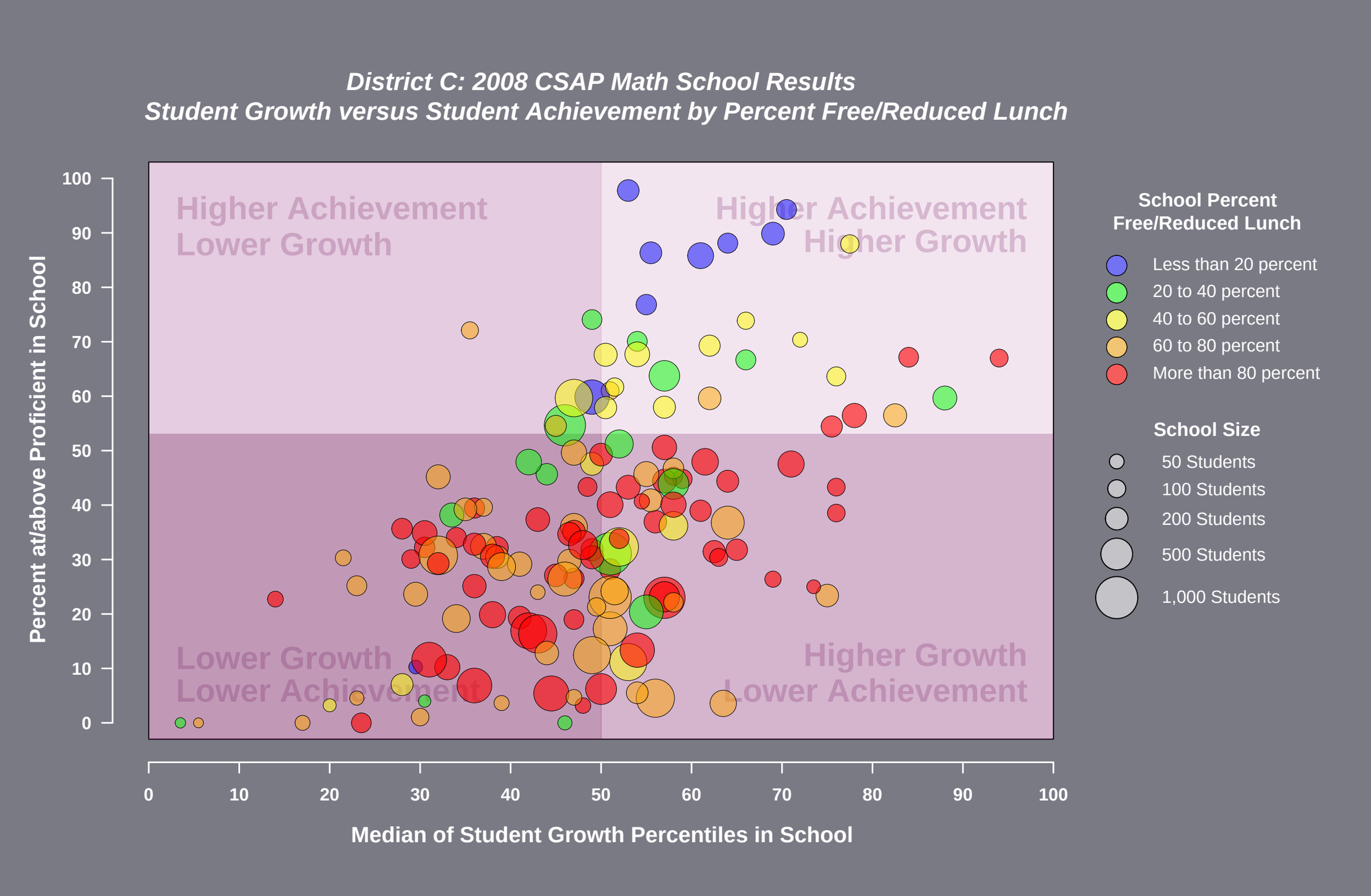
Growth

Level	Percentiles
High	66th – 99th
Typical	35th – 65th
Low	1st – 34th

	Grade 3 2006	Grade 4 2007	Grade 5 2008	Grade 6 2009	Next Year	
Scale Score	462	539	563	609		Achievement
Achievement Level	Unsatisfactory	Part Proficient	Part Proficient	Proficient		
Growth Percentile		66	66	90		Growth
Growth Level		High	High	High		

Going from students to groups of students

- It's of interest to examine schools where students demonstrate, on average, extraordinarily high and low student growth.
- To summarize the student growth percentiles associated with a school (or other grouping) one calculates the median of the student growth percentiles.
- If students were randomly assigned to schools, one expects to see a median of 50.
- Values greatly above or below 50 are of interest in identifying best practices or providing extra support.
- Examining growth with achievement sheds new light on school performance.


District C: 2008 CSAP Math School Results
Student Growth versus Student Achievement by Percent Free/Reduced Lunch
Higher Achievement
Lower Growth
Higher Achievement
Higher Growth
Lower Growth
Lower Achievement
Higher Growth
Lower Achievement
Percent at/above Proficient in School
0
10
20
30
40
50
60
70
80
90
100
Median of Student Growth Percentiles in School
0
10
20
30
40
50
60
70
80
90
100
School Percent Free/Reduced Lunch
Less than 20 percent
20 to 40 percent
40 to 60 percent
60 to 80 percent
More than 80 percent
School Size
50 Students
100 Students
200 Students
500 Students
1,000 Students

Growth, Effectiveness, and Value-Added

Fundamental Premise

"Good" schools bring about student growth in excess of that found at "bad" schools.

- "Good schools" are often called highly effective schools.
- What's the relationship between growth and effectiveness?
- Effectiveness indicates who/what is responsible for the growth (value-added models).
- Yen's questions make no mention of effectiveness placing a greater emphasis on description
- This work group's name reflects a more value-added emphasis.

Next Steps for growth data

- Calculating growth quantities is the first and easiest step in the development and deployment of a growth model.
- The end goal is to change conversations about education through the examination and use of data.
- To this end, Colorado has been inspired by the Web 2.0 user-centered design philosophies and technologies.
- The goal is to produce data visualization and social networking tools to turn data into information for all education stakeholders and ultimately change conversations about education.

Descriptive Accountability

"Accountability system results can have value without making causal inferences about school quality, solely from the results of student achievement measures and demographic characteristics. Treating the results as descriptive information and for identification of schools that require more intensive investigation of organizational and instructional process characteristics are potentially of considerable value. Rather than using the results of the accountability system as the sole determiner of sanctions for schools, they could be used to flag schools that need more intensive investigation to reach sound conclusions about needed improvements or judgments about quality."

R. L. Linn (2008)

Descriptive Accountability

"This is the difference between a retrospective question of identifying fault as opposed to a prospective strategy to engineer some corrective measure, almost independent of considering whether there was blame-worthiness. And to move away from the blame-worthiness paradigm toward something that is more regulatory in nature where one might seize upon disparities or circumstances that are for some reason deemed unacceptable and engineer the interventions needed to bring about the necessary change. . . . It's the no-fault gap closing strategy in which the effort is to build a consensus about a vision of an improved society rather than figure out where's the person we want to pillory."

C. Edley (2006)

Web 2.0: Data Visualization and Social Networking

- The Colorado Department of Education and the Center for Assessment have been working for the last two year on developing next generation data visualization to accompany growth model data.
- The goal: Transform conversations about education through active engagement with data (i.e., evidence).
- Our efforts have received tremendous interest and recognition:
  - Received (just recently) the 2010 annual award for Outstanding Dissemination of Educational Measurement Concepts to the Public by the National Council on Measurement in Education.
  - Recognized by Adobe for innovative uses of their technology as an Adobe Max Award finalist in October, 2009.
  - Multiple states signing MOUs to co-develop a cloud-based reporting platform in a non-proprietary fashion.
  - Colorado recently devoted $2.5 million of stimulus funds to the development efforts

Web 2.0: Data Visualization and Social Networking

With a collaborative spirit, with a collaborative platform where people can upload data, explore data, compare solutions, discuss the results, build consensus, we can engage passionate people, local communities, media and this will raise–incredibly–the amount of people who can understand what is going on.

And this would have fantastic outcomes: the engagement of people, especially new generations; it would increase knowledge, unlock statistics, improve transparency and accountability of public policies, change culture, increase numerary, and in the end, improve democracy and welfare.

E. Giovannini, Chief Statistician, OECD. June 2007

References

Betebenner, D. W. (2008).
Toward a normative understanding of student growth.
In Ryan, K. E. and Shepard, L. A., editors, *The Future of Test-Based Educational Accountability*, pages 155–170. Taylor & Francis, New York.

Betebenner, D. W. (2009).
Norm- and criterion-referenced student growth.
Educational Measurement: Issues and Practice, 28(4):42–51.

Edley, C. (2006).
Educational “Opportunity" is the highest civil rights priority. So what should researchers and lawyers do about it?
Retrieved June 22, 2006 from the World Wide Web: `http://www.softconference.com/MEDIA/WMP/260407/#43.010`.

Linn, R. L. (2008).
Educational accountability systems.
In *The Future of Test-Based Educational Accountability*, pages 3–24. Taylor & Francis, New York.

Yen, W. M. (2007).
Vertical scaling and No Child Left Behind.
In Dorans, N. J., Pommerich, M., and Holland, P. W., editors, *Linking and Aligning Scores and Scales*, pages 273–283. Springer, New York.